British Journal of Educational Technology *Vol 51 No 3 2020* 611–628
doi:10.1111/bjet.12900

The use of augmented reality to foster conceptual knowledge acquisition in STEM laboratory courses—Theoretical background and empirical results

Kristin Altmeyer, Sebastian Kapp, Michael Thees, Sarah Malone, Jochen Kuhn and Roland Brünken

Kristin Altmeyer is a PhD student at the Department of Education, Saarland University, Saarbrücken, Germany. Her research interests include design principles for multimedia learning in maths and physics. Sebastian Kapp is a PhD student at the Physics Education Research Group, Department of Physics, TU Kaiserslautern, Germany. His research focuses on augmented reality learning scenarios for physics experiments. Michael Thees is a PhD student at the Physics Education Research Group, Department of Physics, TU Kaiserslautern, Germany. His research focuses on augmented reality learning scenarios for physics experiments. Sarah Malone is a senior research scientist at the Department of Education at Saarland University, Saarbrücken, Germany. She received her PhD in the field of psychology at Saarland University in 2012. Her research focuses on learning and testing with multiple representations in different educational domains. Jochen Kuhn is a full professor and Head of the Physics Education Research Group at the Department of Physics, TU Kaiserslautern, Germany. He received his PhD in the field of physics in 2002 and his Habilitation in Physics Education in 2009, both at University of Koblenz-Landau. His research focuses on learning with multiple representations in STEM education in general and in physics education in particular using common and advanced multimedia technology (eg, smartphones, Tablet PC, AR-smartglasses). He uses paper-and-pencil instruments as well as physiological methods (especially eyetracking) to analyze learning before, during and after intervention. Roland Brünken is a full professor at the Department of Education, Saarland University, Saarbrücken, Germany. He received his PhD in the field of psychology at Erfurt University in 1998. His research interests cover multimedia learning, learning with multiple representations and cognitive load measurement. Address for correspondence: Sarah Malone, Department of Education, Saarland University, Campus A4 2, 66123 Saarbrücken, Germany. Email: s.malone@mx.uni-saarland.de

Abstract

Learning with hands-on experiments can be supported by providing essential information virtually during lab work. Augmented reality (AR) appears especially suitable for presenting information during experimentation, as it can be used to integrate both physical and virtual lab work. Virtual information can be displayed in close spatial proximity to the correspondent components in the experimentation environment, thereby ensuring a basic design principle for multimedia instruction: the spatial contiguity principle. The latter is assumed to reduce learners' extraneous cognitive load and foster generative processing, which supports conceptual knowledge acquisition. For the present study, a tablet-based AR application has been developed to support learning from hands-on experiments in physics education. Real-time measurement data were displayed directly above the components of electric circuits, which were constructed by the learners during lab work. In a two group pretest–posttest design, we compared university students' (N = 50) perceived cognitive load and conceptual knowledge gain for both the AR-supported and a matching non-AR learning environment. Whereas participants in both conditions gave comparable ratings for cognitive load, learning gains in conceptual knowledge were only detectable for the AR-supported lab work.

Practitioner Notes

What is already known about this topic

- Augmented reality (AR) is used in a variety of educational settings to provide additional virtual information to learners.
- AR tools can boost learning effects compared to non-AR treatments.
- There is a need for research to identify design features that allow students to acquire basic competences related to STEM disciplines.

What this paper adds

- Even the acquisition of conceptual knowledge, which is particularly difficult to address in inquiry-based learning, can be fostered by the use of AR in physical experimentation.
- Basic design principles for multimedia instruction, such as the spatial contiguity principle, can be applied to guide the development of AR-supported learning environments.
- Both applied learning environments do not inhibit the learning process through occupying mental resources by extraneous processing.

Implications for practice and/or policy

- A sophisticated, high-usability AR application for tablet-PCs was developed to provide real-time measurement data during hands-on experiments.
- The tablet-based AR application developed in the present research promises to be adequate in terms of usability in school settings.
- Tablet-based AR can be an affordable alternative to smartglasses to support physical experimentation in schools.

Introduction

According to several literature reviews, the R&D-initiatives of augmented reality (AR) technology in educational scenarios have been quickly taking root, revealing the possibilities and challenges of combining real and virtual elements in various fields of learning and education (Akçayir & Akçayir, 2017; Bacca, Baldiris, Fabregat, Graf, & Kinshuk, 2014; Garzón & Acevedo, 2019; Ibáñez & Delgado-Kloos, 2018; Radu, 2014). In their meta-analysis, Garzón and Acevedo (2019) revealed that AR favors students' learning gains with a medium-sized effect (d = 0.68), and in general, the relevant literature classifies AR as a beneficial tool for education. However, it also points towards the most common challenges researchers face during the development and implementation of such high-tech learning environments: the usability of the whole system and the management of possible cognitive overload situations leading to the question of how best to design such learning environments.

Modern laboratory work, which is a common method of instruction in science teaching, seems to be particularly suited for the application of AR, because the learners are required to interact with physical objects (ie, experimentation materials) on the one hand and virtual objects (ie, measurement data) on the other. In the present study, laboratory learning environments were enhanced with real-time visualizations of measurement data, which were automatically presented close to the corresponding real objects. Thus, AR was used to create contiguity in time and space of essential information in lab work. In general, laboratories allow for unique learning experiences. However, positive learning outcomes are not guaranteed per se (Finkelstein *et al*., 2005; Hofstein & Lunetta, 2004; Holmes & Bonn, 2015; Husnaini & Chen, 2019; Kapici, Akcay, & de Jong, 2019; Volkwyn, Allie, Buffler, & Lubben, 2008; Wieman & Holmes, 2015; Wilcox & Lewandowski, 2017). Especially in the context of STEM (science, technology, engineering and mathematics) education, laboratory experiences, which combine virtual and physical components are claimed to promote learning processes (Jones, 2019; de Jong, Linn, & Zacharia, 2013).

Based on major ideas of cognitive constructivism and multimedia learning theories, such as the principle of contiguity (Mayer, 2009; Mayer & Moreno, 2003), AR can be used to create an integrated format consisting of the physical laboratory and the virtual information. As demonstrated in the present research, this can be realized by presenting real-time measurement data in close spatial proximity to the corresponding experimentation components.

The overarching goals of using AR technology in this manner were to reduce students' extraneous processing and to promote the construction of conceptual knowledge by providing contiguous information (Dounas-Frazer & Lewandowski, 2018; Kuhn *et al*., 2016; Lai, Chen, & Lee, 2019; Strzys, Thees, Kapp, & Kuhn, 2019; Thees *et al*., under review).

The AR-supported condition was compared to a non-AR condition, whereby the real-time measurement data were displayed in a grid on a tablet PC, resulting in a spatial split-source format. The two conditions were contrasted for changes in conceptual knowledge and perceived cognitive load. Due to their broad acceptance and availability, tablet PCs were used as displaying technology instead of smartglasses, which have also proven useful for this type of application (Kuhn *et al*., 2016; Strzys *et al*., 2018, 2019; Thees *et al*., under review).

The content of the laboratory learning experience in the current study dealt with the topic of electricity.

Theoretical and empirical background

Technology-enhanced learning with physical experimentation includes elements of multimedia learning: visual (such as real electric circuits built of cords, resistors and power supply; digital representation of data as needle deflection) and verbal information (worksheets, measurement data) are used together for knowledge construction (eg, guided deduction of rules). According to Santos *et al*. (2014), displaying AR information in particular supports the transformation of traditional learning environments into multimedia settings by literally integrating additional external representations into the physical environment. This perspective allows to apply approved theoretical models on multimedia learning to predict outcomes of AR-learning scenarios.

There are two influential and related theories, which consider learning as information processing and guide research on multimedia instruction: Cognitive load theory (CLT; Sweller, 1988; Van Merriënboer & Sweller, 2005) and the cognitive theory of multimedia learning (CTML; Mayer, 2005, 2009). These theoretical approaches are used to predict and explain the effectiveness of multimedia learning environments by considering memory processes and various design principles for multimedia instruction have been deduced from them.

CLT (Sweller, 1988; Van Merriënboer & Sweller, 2005) is based on assumptions regarding the limited capacity of the working memory in terms of the amount of information that can be processed simultaneously, as well as the amount of time information is maintained for processing. According to research (Leppink & Van der Heuvel, 2015; Sweller, van Merriënboer, & Paas, 1998, 2019), cognitive load is understood as composed of three types of load: (a) intrinsic cognitive load (ICL) refers to the complexity of the information that has to be processed and is therefore determined by the actual task as well as context-specific prior knowledge; (b) extraneous cognitive load (ECL) is assigned to task-irrelevant cognitive processes that occupy working memory resources and can be influenced by the instructional design, eg, how the information is presented; (c) germane cognitive load (GCL) refers to the amount of cognitive resources dedicated to processing information into knowledge structures, ie, the actual learning process. In recent years, researchers have suggested a realignment towards a two-factor ICL/ECL model, in which GCL is an indicator of the effectiveness of the individual learning process (Sweller *et al.*, 2019).

One of the main theoretical assumptions of the CTML, which is based on a view of active knowledge construction, is that learners do not simply absorb new information but that meaningful learning requires them to actively engage in using the provided information to construct mental representations of the learning content (Mayer, Moreno, Boire, & Vagge, 1999). According to the CTML, this active processing of multimedia instruction involves three main processes in working memory: *selection*, *organization* and *integration*. The first process involves selecting relevant visual or verbal elements from the presented multimedia information. Thereafter, the selected information is organized into coherent visual and coherent verbal mental representations which are held in the visual and verbal working memory. Finally, these representations are integrated with both each other and with appropriate prior knowledge stored in long-term memory. However, the CTML also assumes that working memory is limited regarding the maximum amount of information which can be actively processed in this memory structure. To foster constructivist learning in multimedia environments, various design principles have been developed that support the economic use of working memory resources (Mayer *et al.*, 1999). A large part of these design principles is particularly geared towards triggering and facilitating integrative processes across the corresponding visual and verbal representations. According to the CTML, generative learning is fostered, when corresponding verbal and visual representations are maintained in working memory simultaneously. Ainsworth (2006) also stated that a simultaneous presentation of multiple representations encourages the learner to integrate information across the different representations.

This assumption leads to the principles of spatial and temporal contiguity (Mayer, 2009; Mayer & Moreno, 2003), which also aim at reducing extraneous processing by presenting corresponding information from different sources simultaneously and thus avoiding a split-attention effect. The split-attention effect (Mayer & Pilegard, 2014; Schroeder & Cenkci, 2018; Sweller & Chandler, 1994; Sweller *et al.*, 2019) implies that the spatial separation of related nonredundant information increases ECL and therefore inhibits the learning process by occupying mental resources. Equally, temporal separation should also be set to a minimum, to reduce the need to maintain a mental representation over a longer period, ie, representational holding (Mayer & Moreno, 2003). A recent review by Schroeder and Cenkci (2018) consolidated the role of the spatial contiguity principle to foster learning processes using integrated design formats instead of split-source formats in different multimedia instructional scenarios.

Purpose of the study

Billinghurst and Dünser (2012) outlined several goals that AR may achieve when applied within educational settings, such as illustrating spatial and temporal concepts, and emphasizing the contextual relationships between real and virtual objects. Prior work showed that AR-based learning environments can avoid split-attention effects in the context of university STEM laboratory courses by presenting virtual information in close spatial proximity to their physical counterparts in real time (Strzys *et al.*, 2018, 2019; Thees *et al.*, under review). In this research, smartglasses were used as an advanced display technology. Due to the uncommonness of this technology and the specific experimental requirements, participants were limited in their interaction with laboratory equipment, which might have hampered the workflow of the experiment. Consequently, it was suggested that subsequent investigations should focus on more common display technologies and contexts that allow for a higher degree of interaction.

Subsequently, the present study was conducted to investigate the use of tablet-based AR to support students' knowledge acquisition concerning changes in electric current and voltage in parallel and series electric circuits. The overall goal was to minimize learners' extraneous processing when integrating physical and virtual information and to support generative learning processes during the conduction of a science experiment. In typical laboratory courses, learners are likely to face the presented information in a split-source format: scientific phenomena have to be observed over a period of time, whereas measurement data have to be retrieved from several devices with their own displays.

The main research aim of the present study was to investigate, whether this spatial and temporal gap between the observation of phenomena and the processing of measurement data, which is assumed to exist in non-AR-supported laboratory learning environments can be closed by providing AR content during lab work. This is in accordance with the possibilities of AR described by Bujak *et al*. (2013) and Radu (2014), and was realized by displaying the measurement data in real-time adjacent to the corresponding physical objects, in order to achieve an integrated format of content-related information sources.

Hypotheses

The tablet-based application was designed to manage participants' cognitive load by combining real and virtual elements in AR, as well as fostering their conceptual learning by visualizing appropriate and complementary multiple representations.

According to the split-attention effect, retrieving related information from separate external sources causes ECL in the learners, whereas the integration of corresponding information sources in time and space, such as real-world objects and (virtual) external representations, reduces those cognitive processes which cause ECL.

H1: Compared to the non-AR presentation of measurement data, tablet-based AR lab work leads to less ECL.

According to the contiguity effect, the simultaneous presentation of visual and verbal information in multimedia learning can support generative processing and thus foster the effectivity of learning, which is defined here as an increase in conceptual knowledge (Pundak & Rozner, 2007; Vosniadou, 2007). The aim of the present research was to investigate whether this design effect also applies to AR-supported lab work.

H2: Tablet-based AR-supported lab work leads to higher learning gains in general topic-related conceptual knowledge than non-AR lab work.

Materials and methods

Development of the AR/non-AR tablet-supported science learning environments

For the present study a universal science learning environment was developed to foster the acquisition of conceptual knowledge regarding electrical circuits. The learning environment consists of custom designed experimentation components with integrated measurement nodes which are able to communicate their measurements to the accompanying mobile application in real time. Tablet-PCs are used to display the real-time measurement data during lab work.

The experimentation components are based on typical modules used in learning scenarios and present the learner with the electronic symbol of the component as well as two sockets to integrate the component into their circuits. In addition, a custom-designed measurement node is integrated which enables the component to constantly measure the applied electric current and voltage. The acquired measurements are then made available using a Bluetooth Low Energy service to be received with the accompanying applications. To make this process transparent to the learner every component has a transparent bottom side. An identical measurement node is also added to the power supply of the experiment using a custom, 3d printed enclosure. Using the Unity3D game development environment, two custom Apple iPad applications were developed. Unity is a cross-platform game engine developed by Unity Technologies. The engine can be used to create three-dimensional, two-dimensional, virtual reality and AR games, as well as simulations and other experiences.

The first application uses a two-dimensional visualization of the measurement data, presenting the visualizations in a grid and therefore spatially split from the experimental components (Figures 1 and 2).

The second application displays the live image of the tablet camera in which the visualizations are anchored to the corresponding experimentation components, resulting in an AR view (Figure 3). The location of the experimentation components in the camera image is identified using visual markers. These markers are recognized using the Vuforia Engine integrated into Unity3D. This view produces a spatially coherent presentation of the measurement data (Kapp *et al.*, 2019). Vuforia is an AR software development kit for mobile devices. It uses computer vision technology to recognize and track planar images and 3D objects in real time. This image registration

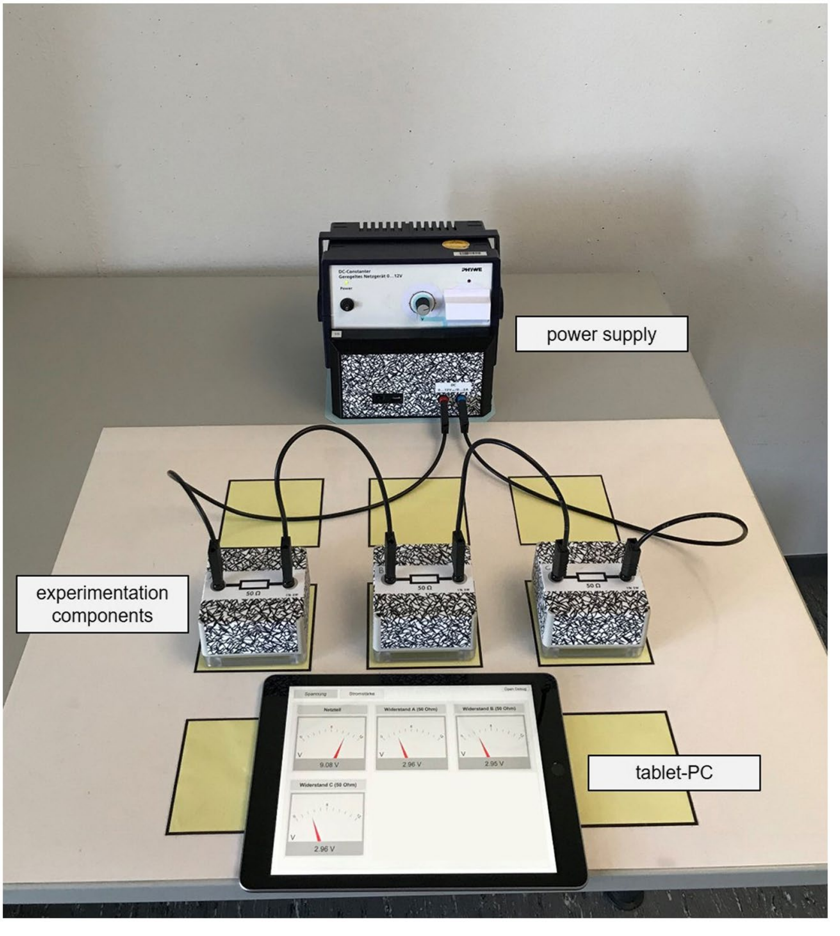


Figure 1: Picture of the experimental setup in the non-AR condition with the tablet PC as the measurement device [Colour figure can be viewed at wileyonlinelibrary.com]

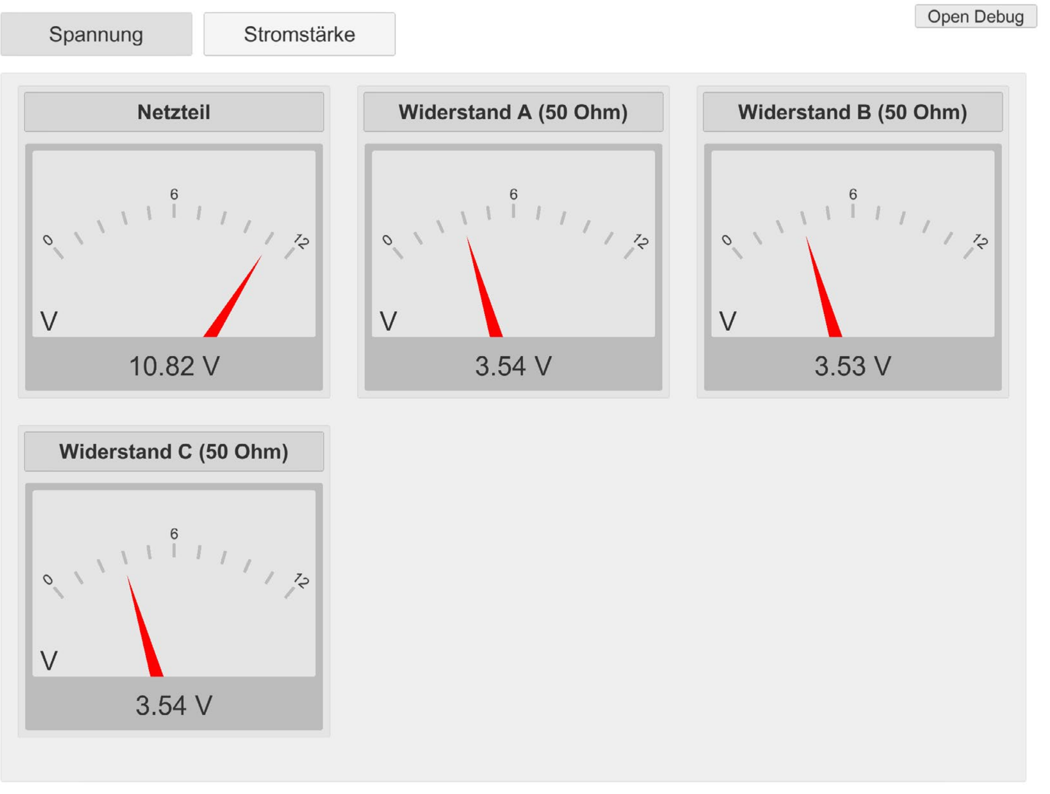


Figure 2: Close-up of the representation of measurement data in the non-AR condition (values and needle deflection) [Colour figure can be viewed at wileyonlinelibrary.com]

capability enables developers to position and orient virtual objects, such as 3D models and other media, in relation to real-world objects when they are viewed through the camera of a mobile device. Thus, it seems as if the virtual objects were part of the real-world scene.

As the applications are set up generically and support a multitude of experiments, an initial configuration of the experiment is offered. Instructors can specify all modules used and configure the connection between the individual measurement nodes and the iPad. Using this configuration, the application then automatically connects to all configured measurement nodes and visualizes their measurement values offering an overview of the whole circuit while enabling the user to switch between voltage and amperage display. In addition, it is possible to hide the visualization of components which are not currently integrated into the circuit. In the present study this option was used by the instructor.

Experimental design and analyses

For the present study, a two group pretest–posttest design was applied. Participants were randomly assigned either to the AR-supported condition or the non-AR condition. Prior conceptual knowledge was pretested in both groups with a concept test on parallel and serial circuits. The same test was used as a posttest after the learning session to evaluate conceptual knowledge gains. Knowledge transfer was measured by a transfer test, which was incorporated in the posttest. Cognitive load was measured by means of a subjective rating scale which was provided to the learners immediately after the lab work.

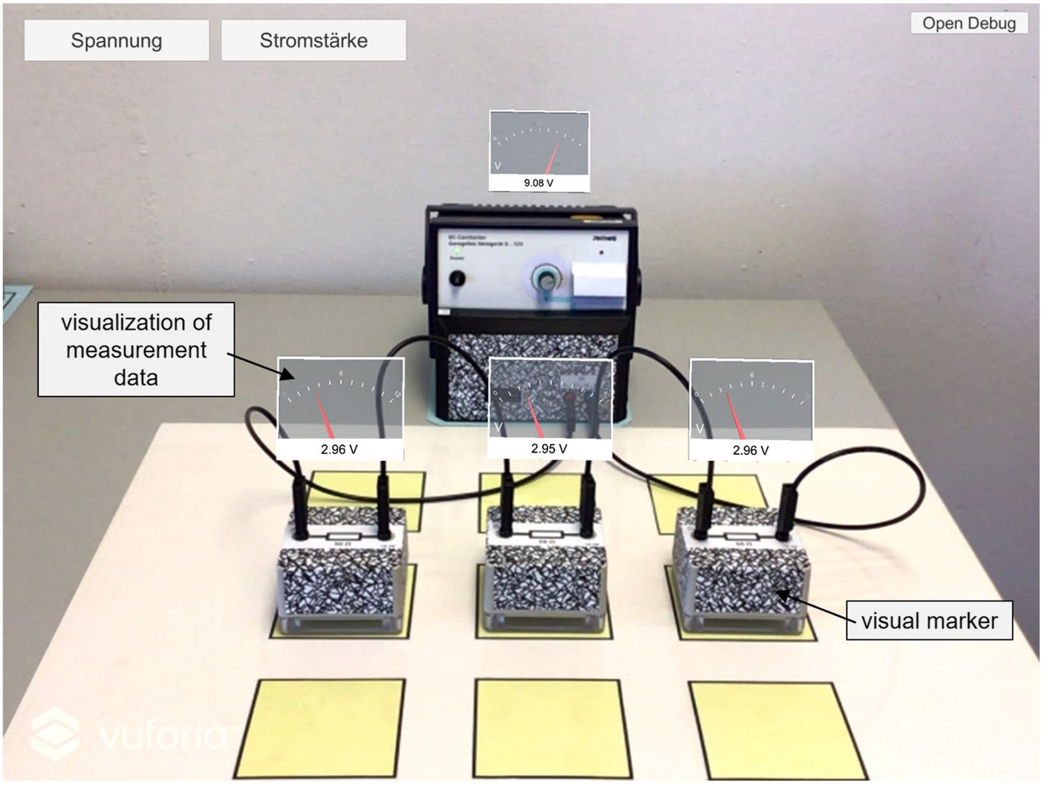


Figure 3: Picture of the experimental setup with AR-view through tablet PC [Colour figure can be viewed at wileyonlinelibrary.com]

The developed AR- and non-AR tablet applications differed slightly in their handling. Therefore, two aspects of usability were measured and compared for the two applications. During the experimentation stage, time on task was logged as an indicator for usability. In addition, usability was measured by means of a subjective rating scale in the posttest.

Procedure and materials

Procedure

The experimental procedure is depicted in Figure 4. First, subjects read general information regarding the study and data protection. Having given written consent, the participants were presented with a short instructional video on a laptop, explaining the basics of voltage, amperage and electric circuits, the aim being to activate the learners' prior knowledge. Subsequently, subjects completed the concept test on parallel and serial electrical circuits for the first time. Upon completing this pretest, participants began the lab work. At this point, participants were assigned either to the AR-supported condition or to the non-AR condition. Subjects of both groups were provided with a tablet PC and familiarized with their workplace, consisting of a table upon which were all the necessary materials to build electric circuits (power supply, cords and resistors) and a work booklet. Using the work booklet, participants were led through three coordinated science experiments dealing with serial and another three experiments dealing with parallel circuits. Half of the sample began with serial circuits, while the other half experimented with parallel circuits first. Each experiment began with the subjects building up an electrical circuit which was

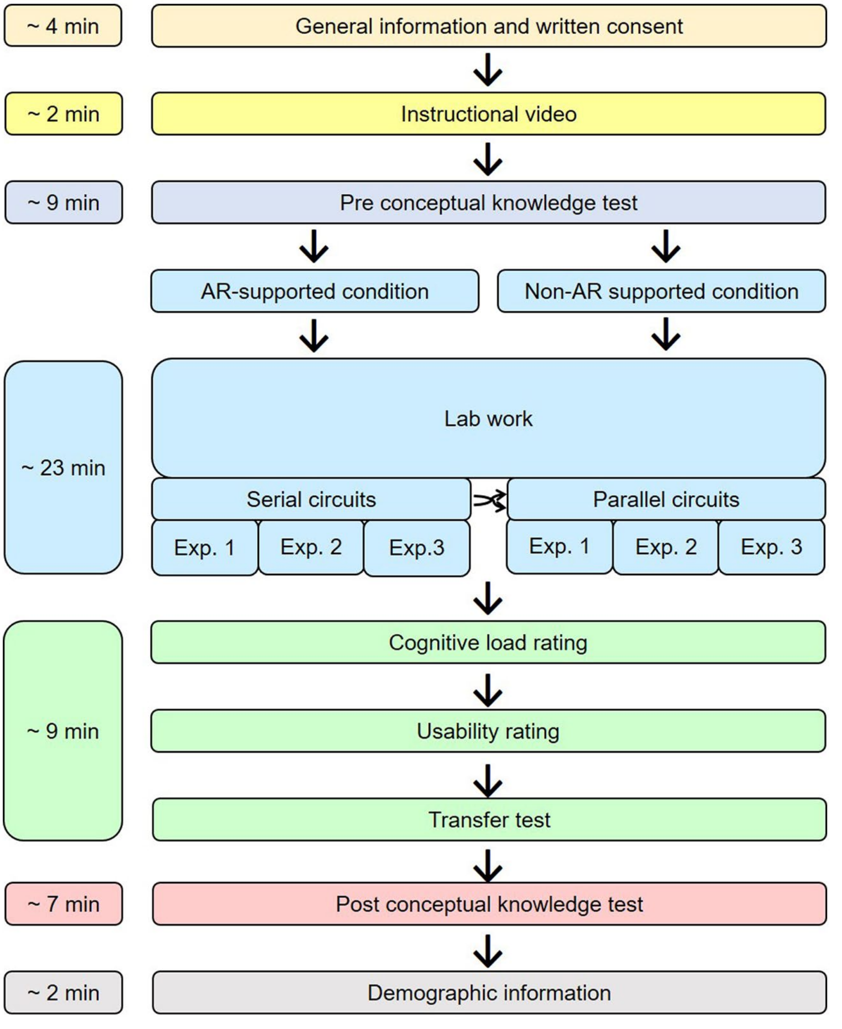


Figure 4: Experimental procedure
[Colour figure can be viewed at wileyonlinelibrary.com]

then photographed by the examiner and corrected if necessary. Subjects were then instructed to apply a potential on the circuits and manipulate the voltage by turning a knob of the power supply. Real-time measurement data were visualized according to the respective condition. Subjects could switch between the display of voltage and amperage by means of a button on the upper left corner of the tablet (see Figure 2; voltage: Spannung; amperage: Stromstärke). During experimentation, the participants answered a set of multiple choice items in the work booklet concerning the relation of voltage or amperage at the electronic components in the current circuit. The time taken for circuit construction as well as to complete the questions in the work booklet was recorded. Subsequent to the lab work phase, participants completed a cognitive load questionnaire, as well as a usability questionnaire, a transfer test and the concept test for a second time. Finally, subjects were asked to provide demographic information (sex, age, course of study), to report their experience with the use of smartphones, tablet PCs and AR applications, and to provide information about their recent school grades (Physics, Maths and German).

Conceptual knowledge

To investigate differences in conceptual knowledge, a power test was used during the pre- and posttest. The test consisted of a selection of 11 items from the concept test evaluating the understanding of electrical circuits developed by Urban-Woldron and Hopf (2012; Cronbachs $\alpha = 0.84$), and was extended by two items of Burde (2018). To mitigate issues due to unfamiliarity with changing terminology and representational forms, the selected items were adapted to match the instructional material of the present study. In addition, two items describing serial circuits were duplicated and adapted to cover parallel circuits.

In total, the conceptual knowledge test consisted of 13 items, 8 of them dealing with serial circuits and 5 with parallel circuits. In the general introduction of the test, relevant terms and symbols were explained. Each single item was composed of an item stem, in which a specific circuit was described (see Table 1). For every item a symbolic representation of the respective electrical circuit was also provided, including the standard symbols for the different components (ie, cords, resistors, lamps, switches). Below the introduction of a particular circuit, the participants were supposed to respond to a question regarding the voltage or the amperage in the respective circuit by choosing the right answer out of several alternatives.

To solve the items of the conceptual knowledge test, knowledge on the respective physical laws concerning amperage and voltage in serial and parallel circuits was required as well as a correct mental representation of the flow of electricity.

Work booklet

To guide the participants through the physical experiments a work booklet was designed. On the first pages, the relevant terminology, symbols and specific representations were introduced, followed by a short statement on measurement accuracy and a visual overview of the components to be used in lab work. After this general introduction, the two tripartite series of experiments (three for each type of circuit) were instructed. The tasks for both circuit types were identical with the only difference being the circuit itself. They began with a short introduction describing the setup of a serial or parallel circuit, accompanied by a circuit diagram and a real-world picture, followed by a set of six multiple choice items. The items were split in three questions regarding the voltage in the circuit and three questions regarding the amperage. All questions were single choice and required qualitative observations of the relation between values. The experiments themselves were structured in the following order: first, a simple circuit containing two identical

Table 1: Conceptual knowledge item (translated and adapted by the authors from Urban-Woldron & Hopf, 2012)

Item 26 *Consider the circuit on the right in which the same voltage is applied across both lamps L_1 and L_2.*

What happens to the voltage across the lamps L_1 and L_2 when the resistance R is increased?

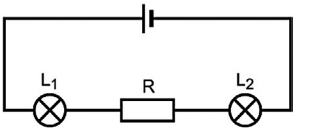
The voltage across L_1 remains the same but the voltage across L_2 is smaller than before.

The voltage across L_1 is smaller than before but the voltage across L_2 remains the same.

The voltages across L_1 and L_2 are larger than before.

The voltages across L_1 and L_2 are smaller than before.

The voltages across L_1 and L_2 remain the same.



resistors was constructed to identify the relation of the voltage and amperage values between them. The second experiment added a third identical resistor to the circuit to validate the relations identified with two resistors. Adding the resistance as factor in the relations, the third experiment used three resistors with different resistances.

A short summary task concluded the booklet, which asked the participants to summarize their observations.

Transfer

The transfer test consisted of 12 multiple choice items. In the item stems, specific electric circuits were described regarding the type of circuit (six parallel, six serial circuits) and the type of integrated resistors (number of resistors; same or different resistance). For each described circuit, a set of five tables containing possible measurements (amperage or voltage) for each component of the circuit and the resulting total amperage resp. voltage was presented. The subjects were instructed to choose the only table containing measurement data which could possibly result from the given circuit.

Cognitive load

To compare the two groups concerning their intrinsic, extraneous and germane processing while performing the experiment, we used an adapted version of the cognitive load scale by Leppink, Paas, Van Gog, Van der Vleuten, and Van Merrienboer (2014) during the posttest. This questionnaire follows the three-factor interpretation of the CLT and is considered to be capable of distinguishing participants' subjective impression of the three sub dimensions of cognitive load (Leppink *et al.*, 2014; Leppink, Paas, Van der Vleuten, Van Gog, & Van Merriënboer, 2013) and can be transferred to other instructional scenarios. We translated the scale into German and adapted it to the context of physics laboratory courses according to the content and structure of the original items. To cover all aspects of laboratory work, such as manipulation of the experimental setup and measurement process, three items were added to the original scale, resulting in 13 items (5 ICL items, 4 ECL items and 4 GCL items). Participants were instructed to indicate their agreement to the presented statements by using a 6-point Likert scale. In a preliminary study (Thees *et al.*, under revision), an exploratory analysis was performed to reveal the internal structure of the adapted instrument (Cronbachs Alpha: ICL = 0.70; ECL = 0.66; GCL = 0.77), which was comparable with the findings from Leppink *et al.* (2014).

Usability

Usability was measured by the System Usability Scale (SUS; Brooke, 1996) translated into German. The subjects were instructed to rate their agreement regarding ten statements on various aspects of system usability. To score the SUS, the values (0–4) for all 10 items were totaled, while taking the inverted items into account. The result was then multiplied by the factor 2.5 (maximum score 100). Overall usability of a specific technology application is represented by the mean usability score across all participants. This mean usability score can be classified according to Bangor, Kortum, and Miller (2009).

Sample

All participants were German university students, randomly assigned to the AR-based ($N = 25$; 20% male; age: $M = 26.71$; $SD = 5.57$) and non-AR-based ($N = 25$; 20% male; age: $M = 25.27$; $SD = 3.59$) tablet environment. Students whose courses of studies were associated with electric circuits (eg, Physics, Systems Engineering) were excluded from the study to avoid excessive domain-specific knowledge.

Since all of the participants were students, the sample of the current study is characterized by young participants that reported considerable familiarity with smartphones and tablets: 90% of them had used a smartphone daily and 74% had used a tablet at least occasionally during the last year. In contrast, the use of AR applications on smartphones or tablets was not very common in the sample of the current study: 82% reported not using them at all.

Results

Descriptive results for cognitive load, pre- and posttest performance in conceptual knowledge, performance in the transfer test, usability and performance in the work booklet items are displayed in Table 2, separated for the two conditions.

Effects on cognitive load

Internal consistencies were mostly acceptable for the scale adapted from Leppink *et al*. (2014) (Cronbach's Alpha: ICL = 0.73; ECL = 0.38; GCL = 0.75).

To contrast the AR-supported with the non-AR tablet environment, for each type of load an independent samples t-test was conducted. No significant differences were detected for ICL, $t(48) < 1$, nor for ECL, $t(48) = 1.62$, $p = 0.112$, or GCL, $t(48) < 1$.

Effects on performance

The first performance variables in focus were the number of correctly solved items and time on task in the concept test. Concerning these variables, the MANOVA revealed no pretest differences in prior conceptual knowledge between the participants of the AR-supported and non-AR conditions, $F(2, 47) < 1$.

A mixed design examining the between-subjects factor treatment (AR-supported vs. non-AR) and the within-subject factor time (pretest vs. posttest) was conducted. The dependent variable was the number of correctly solved items in the concept test. No significant main effects for time, $F(1, 48) < 1$ or treatment, $F(1, 48) < 1$ were found. The time-by-treatment interaction (one-tailed)

Table 2: Means (M) and standard deviations (SD) for dependent variables, separate for the AR and non-AR condition

	AR-supported lab work	*Non-AR lab work*
AR-supported lab work	*M (SD)*	*M (SD)*
Variable		
Cognitive load (range 1–7)		
ICL	1.49 (0.56)	1.66 (0.68)
ECL	1.31 (0.49)	1.54 (0.51)
GCL	4.45 (1.05)	4.42 (0.93)
Pre-conceptual knowledge (max 13)	6.12 (1.99)	6.44 (1.94)
Serial Circuits	3.96 (1.43)	4.24 (1.71)
Parallel Circuits	2.16 (1.25)	2.20 (1.15)
Time (s)	537.00 (141.00)	548.00 (155.00)
Post-conceptual knowledge (max 13)	6.96 (2.11)	6.16 (1.86)
Serial Circuits	3.96 (1.57)	3.48 (1.36)
Parallel Circuits	3.00 (1.19)	2.68 (1.11)
Time (s)	401.00 (162.00)	397.00 (164.00)
Transfer (max 12)	7.72 (3.86)	7.56 (4.19)
Time (s)	290.00 (76.00)	338.00 (164.00)
System usability (max 100)	88.40 (10.36)	89.20 (14.36)
Work booklet item score (max 36)	33.96 (3.14)	33.88 (5.49)
Work booklet time (s)	1010.00 (259.08)	967.96 (406.51)

was significant, $F(1, 48) = 3.48$, $p = 0.034$, $\eta_p^2 = 0.068$, indicating that AR had an impact on the learning gain from pre- to posttest in the conceptual knowledge items. Whereas the subjects in the AR-supported learning condition on average solved 0.84 ($SD = 1.84$) items more in the post- than in the pretest, performance in the non-AR condition even slightly declined in the posttest ($M = -0.28$, $SD = 2.37$). To further specify the learning gains, descriptive results for the concept test are displayed separately (see Table 2) for the items about serial and parallel circuits. The means imply that learning gains in the AR-supported group were only detectable for the concept of parallel circuits. Regarding the non-AR-supported condition, the slight losses in conceptual knowledge were due only to decreasing performance in the items on serial circuits.

Applying the respective mixed design in order to analyze the dependent variable time to complete the concept test revealed no interaction effect, $F(1, 48) < 1$ and no main effect for treatment, $F(1, 48) < 1$. However, a main effect for the factor time was detected, $F(1, 48) = 49.90$, $p < 0.001$, $\eta_p^2 = 0.510$, indicating that participants performed faster in the posttest (see Table 2).

Further performance measures were assumed to indicate possible learning benefits for a particular learning environment where time to complete and the number of correctly solved items in the transfer test. A MANOVA analyzing the overall number of correctly solved items and time on transfer test as dependent variables showed no significant differences between the two tablet environments, $F(2, 47) < 1$.

Usability

The average usability score (computed following Brooke, 1996) was 88.40 ($SD = 10.36$; range: 57.50–100) for the tablet-based AR application, and for the non-AR application 89.20 ($SD = 10.36$; range: 45–100). Both mean usability scores can be classified as "excellent," the second best rating on Bangor *et al*. (2009)'s adjective rating scale. The scores did not differ significantly between the AR-supported and non-AR condition, $F(1, 48) < 1$. Further analyses revealed that the mean usability ratings on the single items (maximum 10) of the SUS ranged between 7.5 (*good*) and 9.6 (*best imaginable*).

However, usability is not only displayed by subjective rating, but can additionally be indicated by process variables, such as time for answering the multiple choice items in the work booklet during the lab work and the number of correctly solved items of the work booklet. Concerning the overall number of correctly solved items and the overall time to complete the items, the MANOVA revealed no differences between the AR-supported and the non-AR tablet environment, $F(2, 47) < 1$.

Discussion and conclusion

The main objective of the present study was to investigate the use of a tablet-based AR application as a tool to support hands-on experimentation with electric circuits. Real-time measurement data were displayed as digits and as virtual needle deflection directly above the corresponding components in electric circuits while the learners manipulated the voltage. This sophisticated application ran smoothly and stable. The participants reported low ICL and ECL, as well as high GCL and excellent usability. Compared to a non-AR-based, spatially delocated, but real-time measures display, the AR condition resulted in slightly higher conceptual knowledge gains. However, learning gains were limited to conceptual knowledge regarding voltage and amperage in parallel circuits.

Table 3 summarizes the main results of the study regarding the two hypotheses.

We used an AR display of measurement data to integrate real and virtual information in physical experimentation settings, in order to reduce spatial split attention effects and provide an

Table 3: Summary of main results regarding hypotheses

Hypothesis (in short)	*Dependent variable*	*Hypothesis supported*
H1 AR-supported lab work leads to less ECL than non-AR lab work.	Subjective ECL rating	No
H2 AR-supported lab work leads to higher learning gains than non-AR lab work.	Conceptual knowledge test score (pre- vs. poststest)	Yes
	Conceptual knowledge test time (pre- vs. posttest)	No
	Transfer test score (post)	No
	Transfer test time (post)	No

integrated instead of a split-source format for learning. H1, which was derived from the assumptions of CLT (Sweller, 1988; Van Merriënboer & Sweller, 2005) and CTML (Mayer, 2005, 2009) and stated that ECL would be reduced in AR-supported lab work, could not be confirmed. This might be due to the fact that subjects in both conditions rated ECL very low, which possibly led to a floor effect that covers possible group differences through restriction of variance. In addition, the only difference between the two conditions was one aspect of spatial contiguity which was fulfilled in the AR-supported but not in the non-AR condition: the spatial contiguity of measurement data and respective physical experimentation components (ie, resistors).

However, further aspects of contiguity were given in both conditions: the spatial and temporal contiguity of the displayed measurement data. Through the AR view as well as on the tablet display in the non-AR condition, the measurements for all included resistors in a specific electric circuit were shown simultaneously and could be perceived altogether at one glance. This is crucial, because by completing the tasks in the work booklet and thereby experiencing the respective physical laws, the learners were required to relate the measurements at the different resistors to one another. Moreover, in both groups the information was presented in real time: at the same time as the participants physically manipulated the voltage, all displayed measurements adjusted to the changes. Therefore, temporal contiguity was fulfilled in both groups.

It is probable that the effect of split attention in the non-AR condition was so small that it did not noticeably impact the learners' mental resources, leading to very low ECL ratings. Thus, from a cognitive load perspective, both learning environments seem to be appropriate to support students' lab work at school. This result is particularly important because in traditional hands-on experiments, split-attention effects can be expected to be much higher. Temporal contiguity is not met there: learners first apply a voltage after which they sequentially measure amperage or voltage at each of the electrical components. That means they have to hold a number of measurements in their working memory and process and integrate the information to infer physical laws.

The second hypothesis could be confirmed as AR-supported learning led to higher conceptual knowledge gains than non-AR learning. Since this result cannot be explained by reduced extraneous processing (H1), the integrated format provided by AR might have triggered more generative processing, leading to enhanced learning processes. However, the advantage of AR learning over non-AR learning was not as pronounced as in previous research (Garzón & Acevedo, 2019; Strzys *et al.*, 2019; Thees *et al.*, under review). This might be because the acquisition of conceptual knowledge can be regarded as a rather ambitious learning objective which is not easily achieved by inquiry-based learning in traditional lab work courses (eg, Kirschner, Sweller & Clark, 2006), but rather in virtual lab work courses (overview: de Jong *et al.*, 2013). However, virtual lab work courses can also have some drawbacks, as no physical equipment is used and so students cannot develop practical laboratory skills, including the troubleshooting of machinery, and they usually do not experience the challenges many scientists face when planning experiments that require careful setup of equipment and observations over long time spans. Consequently, AR-based lab work learning environments could combine the advantages of both physical and virtual lab work settings.

Less distal and more sensitive performance measures would have possibly made learning gains more visible and emphasized the effects of AR. Moreover, the developed AR application could also be used with smartglasses, potentially leading to more pronounced effects compared to tablet-based AR and non-AR conditions (Strzys *et al.*, 2019; Thees *et al.*, under review). Smartglass applications render the learners' hands free for experimentation, which could foster embodied learning (Korbach, Ginns, Brünken, & Park, accepted) and reduce interruptions, caused by tablet use, leading to more temporal contiguity of corresponding information. However, from a practical perspective, the present research indicates that effective AR-supported learning does not necessarily require specific equipment such as smartglasses, but that tablet-PCs or even smartphones can be used. These common technologies have a multitude of advantages. They allow for location independent learning, are easy to access, and are less expensive for the learners themselves or for educational institutions. Furthermore, their general use is familiar to a great proportion of society, and yet there are many AR applications and skilled developers.

In summary, the present research revealed, that a tablet-based AR-application can foster conceptual knowledge acquisition in Physics lab work. However, more research is required to replicate these findings in different learning settings and other instructional domains. Since the present study was conducted in single-subjects experiments with a psychologist as instructor, it remains unclear whether the application is also useful in less structured learning settings. Lab work in school classes is usually guided by teachers and students, normally working together in pairs or small groups. Therefore, subsequent research should test the application in real school settings to establish whether the system is easy to apply and manage by teachers and whether it can be successfully adapted to the demands of classroom learning, such as collaborative learning contexts.

Moreover, subsequent research should investigate whether the results of the current study can be replicated in similar learning environments, such as in other disciplines (ie, Chemistry) that incorporate lab work for learning.

Subsequent research is also needed to reveal whether the benefit of AR regarding the possibility of providing information from different sources contiguously regarding time and space, is generalizable to more distant educational disciplines. One possible field of application is language learning. In that context, learners might, eg, profit from an AR-display of vocabularies close to the respective real-life equivalents.

Future research can also reveal whether the success of the applied AR-supported learning environment is influenced by the individual differences of the learners, such as their cognitive abilities, as well as their emotions (as suggested by the Integrated Cognitive Affective Model of Learning; Plass & Kalyuga, 2019). Furthermore, it would be helpful to gain further insight into the learning processes and the learners' interaction with components and visualizations during the lab work process. In their eye tracking study, Chien, Tsai, Chen, Chang, and Chen (2015) revealed that learners allocated their visual attention differently, when learning in virtual compared to physical laboratories. Since AR-supported lab work integrates both, future studies should consider process-based measures, such as eye tracking to investigate, how AR affects attentional processes during learning.

In conclusion, the newly established AR-based representation of measurement data during lab work has been shown to be appropriate as a tool to support physical experimentation. More research is necessary to prove whether it can be transferred to other learning domains and contexts.

Acknowledgements
The data set this paper draws upon was collected as part of the research project GeAR funded by the Federal Ministry of Education and Research (BMBF).

Statements on open data, ethics and conflict of interest
The relevant data from the current data set can be made available by contacting the corresponding author.

The participating students took part voluntarily and with informed consent. The data used in this study were anonymized.

No conflict of interest exists in the submission of this manuscript.

References
Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001
Akçayir, M., & Akçayir, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, 1–11.
Bacca, J., Baldiris, S., Fabregat, R., Graf, S., & Kinshuk. (2014). Augmented reality trends in education: A systematic review of research and applications. *Educational Technology & Society*, *17*(4), 133–149.
Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, *4*(3), 114–123.
Billinghurst, M., & Dünser, A. (2012). Augmented reality in the classroom. *Computer*, *45*(7), 56–63.
Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). London, UK: Taylor & Francis.
Bujak, K. R., Radu, I., Catrambone, R., MacIntyre, B., Zheng, R., & Golubski, G. (2013). A psychological perspective on augmented reality in the mathematics classroom. *Computers & Education*, *68*, 536–544.
Burde, J.-P. (2018). *Konzeption und Evaluation eines Unterrichtskonzepts zu einfachen Stromkreisen auf Basis des Elektronengasmodells*. Berlin: Logos Verlag Berlin. https://doi.org/10.30819/4726
Chien, K.-P., Tsai, C.-Y., Chen, H.-L., Chang, W.-H., & Chen, S. (2015). Learning differences and eye fixation patterns in virtual and physical science laboratories. *Computers & Education*, *82*, 191–201. https://doi.org/10.1016/j.compedu.2014.11.023.
de Jong, T., Linn, M., & Zacharia, Z. (2013). Physical and virtual laboratories in science and engineering education. *Science*, *340*(6130), 305–308.
Dounas-Frazer, D. R., & Lewandowski, H. J. (2018). The modelling framework for experimental physics: Description, development, and applications. *European Journal of Physics*, *39*, 064005.
Finkelstein, N., Adams, W., Keller, C., Kohl, P., Perkins, K., Podolefsky, N., … LeMaster, R. (2005). When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics Physics Education Research*, *1*(1), 010103.
Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of Augmented Reality on students' learning gains. *Educational Research Review*, *27*, 244–260. https://doi.org/10.1016/j.edurev.2019.04.001
Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*(1), 28–54.

Holmes, N. G., & Bonn, D. A. (2015). Quantitative comparisons to promote inquiry in the introductory physics lab. *The Physics Teacher*, *53*, 352–355.
Husnaini, S., & Chen, S. (2019). Effects of guided inquiry virtual and physical laboratories on conceptual understanding, inquiry performance, scientific inquiry self-efficacy, and enjoyment. *Physical Review Physics Education Research*, *15*(1), 010119.
Ibáñez, M.-B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education*, *123*, 109–123.
Jones, N. (2019). The virtual lab: Can a simulated laboratory experience provide the same benefits for students as access to a real-world lab? *Nature*, *507*, 5–7.
Kapici, H., Akcay, H., & de Jong, T. (2019). Using hands-on and virtual laboratories alone or together—Which works better for acquiring knowledge and skills? *Journal of Science Education and Technology*, *28*(3), 231–250.
Kapp, S., Thees, M., Strzys, M. P., Beil, F., Kuhn, J., Amiraslanov, O., … Wehn, N. (2019). Augmenting Kirchhoff's laws: Using augmented reality and smartglasses to enhance conceptual electrical experiments for high school students. *The Physics Teacher*, *57*(1), 52–53.
Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86.
Korbach, A., Ginns, P., Brünken, R., & Park, B. (2019). Should learners use their hands for learning? Results from an eye-tracking study. *Journal of Computer Assisted Learning*, 1–12. https://doi.org/10.1111/jcal.12396
Kuhn, J., Lukowicz, P., Hirth, M., Poxrucker, A., Weppner, J., & Younas, J. (2016). gPhysics—Using smart glasses for head-centered, context-aware learning in physics experiments. *IEEE Transactions on Learning Technologies*, *9*(4), 304–317.
Lai, A., Chen, C., & Lee, G. (2019). An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *British Journal of Educational Technology*, *50*(1), 232–247.
Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1
Leppink, J., Paas, F., Van Gog, T., Van der Vleuten, C. P. M., & Van Merrienboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, *30*, 32–42.
Leppink, J., & Van der Heuvel, A. (2015). The evolution of cognitive load theory and its application to medical education. *Perspectives on Medical Education*, *4*, 119–127.
Mayer, R. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. Cambridge, MA: Cambridge University Press.
Mayer, R. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.
Mayer, R., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 43–52.
Mayer, R. E., Moreno, R., Boire, M., & Vagge, S. (1999). Maximizing constructivist learning from multimedia communications by minimizing cognitive load. *Journal of Educational Psychology*, *91*(4), 638–643. https://doi.org/10.1037/0022-0663.91.4.638
Mayer, R., & Pilegard, C. (2014).Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality principles. In R. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 316–344). New York, NY: Cambridge University Press.
Plass, J. L., & Kalyuga, S. (2019). Four ways of considering emotion in cognitive load theory. *Educational Psychology Review*, *31*(2), 339–359.
Pundak, D., & Rozner, S. (2007). Empowering engineering college staff to adopt active learning methods. *Journal of Science Education and Technology*, *17*(2), 152–163.
Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing*, *18*(6), 1533–1543.

Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., & Kato, H. (2014). Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Transactions on Learning*, *7*(1), 38–56.

Schroeder, N. L., & Cenkci, A. T. (2018). Spatial contiguity and spatial split-attention effects in multimedia learning environments: A meta-analysis. *Educational Psychology Review*, *30*, 679–701.

Strzys, M. P., Kapp, S., Thees, M., Klein, P., Lukowicz, P., Knierim, P., Schmidt, A., & Kuhn, J. (2018). Physics holo.lab learning experience: Using Smartglasses for Augmented Reality labwork to foster the concepts of heat conduction. *European Journal of Physics*, *39*(3), 035703.

Strzys, M. P., Thees, M., Kapp, S., & Kuhn, J. (2019).Smartglasses in STEM laboratory courses—The augmented thermal flux experiment. In A. Traxler, Y. Cao, & S. Wolf (Eds.), *2018 PERC Proceedings [Washington, DC, August 01–03, 2018]*. College Park, MD: AAPT. https://doi.org/10.1119/perc.2018.pr.Strzys

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257–285.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, *12*(3), 185–233.

Sweller, J., van Merriënboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2019). Using Smartglasses as an efficient augmented reality tool for enhancing university STEM laboratory courses—Theoretical derivations and empirical findings from a multimedia learning perspective. *Computers in Human Behavior 102* (2020), under review.

Urban-Woldron, H., & Hopf, M. (2012). Entwicklung eines Testinstruments zum Verständnis in der Elektrizitätslehre [Development of a diagnostic instrument for testing student understanding of basic electricity concepts]. *Zeitschrift für Didaktik der Naturwissenschaften*, *18*(2012), 201–227.

Van Merriënboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177.

Volkwyn, T. S., Allie, S., Buffler, A., & Lubben, F. (2008). Impact of a conventional introductory laboratory course on the understanding of measurement. *Physical Review Special Topics—Physics Education Research*, *4*(1), 010108.

Vosniadou, S. (2007). Conceptual change and education. *Human Development*, *50*, 47–54.

Wieman, C., & Holmes, N. G. (2015). Measuring the impact of an instructional laboratory on the learning of introductory physics. *American Journal of Physics*, *83*(11), 972–978.

Wilcox, B. R., & Lewandowski, H. J. (2017). Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics. *Physical Review Physics Education Research*, *13*(1), 010108.